The Thesis Committee for Wiley Charles Jennings
Certifies that this is the approved version of the following thesis:

Indoor biological exposures: what can HVAC filter dust tell us?

APPROVED BY
SUPERVISING COMMITTEE:

Supervisor: _____
Kerry A. Kinney

_____
Juan Pedro Maestre Wic

Indoor biological exposures: what can HVAC filter dust tell us?

by

Wiley Charles Jennings, B.A.

Thesis

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Master of Science in Engineering

The University of Texas at Austin

May 2016

Acknowledgements

The work presented in this thesis is an important piece of a much larger study investigating indoor biological exposures and childhood asthma, which is in turn one of three current projects in the lab of Dr. Kerry Kinney investigating indoor exposures and childhood wellbeing and academic performance. I would like to thank my advisor, Dr. Kerry Kinney, for support and guidance during my Masters research. I would also like to thank her for undertaking projects that are multidisciplinary, which is crucial for advancing understanding at the intersection of public health and environmental science and engineering. I would also like to thank Dr. Juan Pedro Maestre for accelerating me my up the steep learning curve that engineers face when entering the arena of microbial community analysis, and for lots of laughter during this process. Finally, I would like to thank Dr. Dennis Wylie for his guidance on statistical methods and R scripting.

Abstract

Indoor biological exposures: what can HVAC filter dust tell us?

Wiley Charles Jennings, M.S.E.

The University of Texas at Austin, 2016

Supervisor: Kerry A. Kinney

Because people in the US spend an estimated 80-90% of their time indoors, much of it at home, understanding the potential health impacts of biological exposures that occur in the home is crucial. Recently, rapid advances in high-throughput DNA sequencing technology have spurred increased study of the relationships between the human and built environment microbiomes. HVAC filters hold promise as long-term, spatially integrated, high volume samplers to characterize the airborne home microbiome. In order to optimize HVAC sampling protocols and improve comparability between studies employing HVAC filters for bacterial community analysis, three HVAC filter dust sampling methods were compared. These three methods, vacuuming the filter surface, swabbing the filter surface, and eluting filter dust in a buffer, were selected as representative of previously published methods. Our findings suggest that vacuum and swab samples produced more repeatable and representative bacterial communities than did elution. Furthermore, given the reduced labor and cost of vacuum and swab methods, and the additional advantage that these two methods may also be applied to sampling dust

from other home surfaces, vacuum and swab sampling of HVAC filter dust are found to be superior to elution.

Table of Contents

List of Tables

List of Figures

1. Introduction

People in the US spend an estimated 80-90% of their time indoors (Adgate et al., 2004; Franck et al., 2003; KLEPEIS et al., 2001), much of it at home. Because of this, the potential health impacts of chemical and biological exposures that occur in the home have been studied for decades (Andelman, 1985; Berglund et al., 1992; Bernstein et al., 2008; Fung and Hughson, 2003; Hardin et al., 2003; Jones, 1999; Lax et al., 2015b; Spengler and Sexton, 1983; Stolwijk, 1992; Wu et al., 2007). Recently, rapid advances in high-throughput DNA sequencing technology (Bartram et al., 2011; Margulies et al., 2005; Quail et al., 2012; Shendure and Ji, 2008; Shendure et al., 2005) have spurred increased study of the relationships between the human and built environment microbiomes, largely bacterial and fungal (Adams et al., 2013, 2014; Barberán et al., 2015a, 2015b; Corsi et al., 2012; Dunn et al., 2013; Fierer et al., 2010; Flores et al., 2013; Kelley and Gilbert, 2013; Kembel et al., 2012, 2014; Lax et al., 2014, 2015b; Luongo et al., 2016; Meadow et al., 2014a, 2014b; Peccia et al., 2011). In particular, several molecular-based studies have linked indoor microbial community exposures to human health outcomes, such as respiratory allergen challenge (Fujimura et al., 2014), and childhood asthma development (Dannemiller et al., 2014; Ege et al., 2011) and severity (Dannemiller et al., 2016).

The most common method used to delineate microbial communities in a home is to collect settled dust samples from dust reservoirs, such as door trims (Barberán et al., 2015b; Dunn et al., 2013), chairs (Dannemiller et al., 2016), floors (Dannemiller et al., 2014), and mattresses (Ege et al., 2011). Alternatives to sampling settled dust on home surfaces include using plastic petri dishes and other passive materials to capture airborne particles as they settle (Adams et al., 2014, 2015; Luongo et al., 2016), as well as small electrokinetic particle samplers (Gordon et al., 2015). These alternative techniques are

advantageous because collection time may be known and collected particles were necessarily airborne. However, since these methods collect relatively low biomass, they are not viable for some downstream analyses.

Recently, HVAC filters have been successfully used to characterize the microbiome of indoor environments (Emerson et al., 2015; Hoisington et al., 2015; Luongo et al., 2016; Noris et al., 2011). This filter forensics approach uses the HVAC filters installed in homes as integrated, long-term samplers of particle-bound contaminants, such as microorganisms (Hoisington et al., 2014) and allergens (Barnes et al., 2015). The potential advantages of using HVAC filter dust are: a) most central HVAC systems have a filter, b) the filters are in place for long, potentially known periods of time, and c) they collect particles from a wide spatial area, acting in essence as a high volume air sampler (Noris et al., 2011). When combined with HVAC system characterization, HVAC filter dust sampling offers a controlled way of detecting and assessing indoor air contaminants present at low concentrations in homes.

To improve comparability between studies employing HVAC filters for microbial community analysis, it is important to compare results produced by difference methods of HVAC filter dust sampling. Understanding these differences is also important for understanding the extent to which dust from any surface sampled with different techniques can be legitimately compared. Several methods of sampling HVAC filters have been previously published (though all not necessarily used for microbial community analysis). Noris et al., 2011 and Hoisington et al., 2014 used a liquid-based extraction involving sonication, vortexing, and filtering to extract dust samples from HVAC filters. Emerson et al., 2015 used swabs to sample the microbial communities on the surface of the HVAC filters and compared them to those captured on sterile petri dishes. Checinska et al. 2015 employed two techniques: vacuuming the surface of the filter and then eluting in a buffer; and removing particles with a scalpel and then eluting in a buffer. Barnes et al., 2015 vacuumed dust from the surface of filters.

These techniques may vary in several regards, including repeatability, biomass recovered, and labor intensiveness. Understanding these differences is key because it has been demonstrated that microbial communities (Andrew Hoisington, 2014; Fahlgren et al., 2010) and other health-relevant biological material (Frankel et al., 2012) in indoor environments may vary significantly by sampler type. To our knowledge, however, no study has compared the effect of sampling method on the microbial community recovered from HVAC filters. Additionally, an understanding of the microbial community variability attributable to sampling methodology may inform decisions about sampling techniques employed for other surfaces as well, such as floors, mattresses, and door trims.

Furthermore, few studies examining indoor microbial communities have examined the sampling variability associated with replicate dust samples. Some studies that have done so have reported very low replicate variability (Adams et al., 2015; Fujimura et al., 2012). However, repeatability in HVAC dust has not yet been investigated. In this study, we perform repeated, randomized sampling without replacement of a single dust matrix in order to both compare sampling methods and assess the variability of microbial communities across an HVAC filter.

2. Materials and Methods

2.1. FILTER INSTALLATION AND COLLECTION

This study is part of the Healthy Homes investigation (HUD: TXHHU0023-13), in which 60 households were recruited based on a resident child's asthma status. Briefly, all homes evaluated were located in rural areas of central Texas, and were sampled in both summer and winter to examine potential relationships between asthma severity and indoor microbial and chemical exposures as measured on home HVAC filters and in settled dust. Brand-new HVAC filters (multiple manufacturers, all ASHRAE Standard 52.2-2007 MERV 7-8) were installed in the air handling unit for 30-45 days. The study presented in this paper comprises an in-depth investigation of the HVAC filters installed for 32 days in a home in Buda, TX. These filters were located in the ceiling at a height of 9 ft. Filters were installed and removed by researchers from The University of Texas at Austin, and then transported and stored at 4 ºC until laboratory processing. Low temperature conditions limit the ability of the microorganisms to reproduce on the filter after removal from the HVAC system (Li and Lin, 2001; Lauber et al. , 2010). Sterile techniques were used for handling the filters in the field and lab.

2.2. HVAC FILTER DUST SAMPLING

In this study, three techniques for removing dust from HVAC filters were investigated: a) swabbing the surface of the filter, b) vacuuming the surface of the filter, and c) cutting pieces of the filter, extracting the dust in a buffer solution, and then filtering. In all three cases, five 1x1-inch squares randomly selected from across a single filter were composited and processed for each sample. To generate a random selection, a 1x1 inch grid was superimposed on the HVAC filters, and each grid was assigned a number. A random number generator without replacement was then used to determine the selection of each subsequent square. Each of the three techniques was used to create

seven samples. No part of the filters was sampled more than once. A second filter from the same home was used for positive controls, as described below.

For each swab sample, a single PBS-T-wetted swab (Floq Swab, Copan, xx, USA) was used. Swabs were then transferred directly to bead beating tubes (Mobio, CA, USA) for DNA extraction. For each vacuum sample, a vacuum thimble was inserted into a clean thermoset plastic nozzle (Indoor Biotechnologies, XX, USA) attached to a vacuum cleaner [Genie Voltaire, XX, USA]. The collected dust cake was then transferred from the thimble to bead beating tubes for DNA extraction. For each cut sample, five squares were cut from the filter and transferred to a pre-sterilized phosphate buffer solution (10 g/L NaCl, 0.25g/L KCl, 1.43 g/L Na2HPO4, 0.25 g/L KH2PO4, DNA-free water, pH 7.0) in a sterile 50 mL centrifuge tube (Thermo Fisher Scientific Inc., Waltham MA). The solution was sonicated and vortexed for 10 minutes, and then pre-filtered through a 20 μm pore size [type] filter (Whatman Ltd., Maidstone United Kingdom). The filtered solution was then vacuum-filtered through a 0.2 μm hydrophobic filter (Millipore, Billerica MA).

Negative controls were obtained by processing an unused swab, an unused thimble (thimble plastic was cut out and placed in the bead beating tube), and a new HVAC filter for swab, vacuum, and cut samples, respectively. Positive controls consisted of spraying pure cultures of *E. coli* and *Nigrospora* spp. on a second used HVAC filter collected from the same house. Dust removal then proceeded as described for each sample type.

2.3. DNA EXTRACTION, PCR, AND SEQUENCING

DNA extraction was conducted as described by Noris et al. (2011). Briefly, the swabs, the filter cake from the thimble (in the case of the vacuumed samples), and the 0.2 μm filter (in the case of the cuts), were added along with 100 μL lysozyme (3mg/mL) and 300 μL phenol-chloroform-isoamyl alcohol (25:24:1) to a bead beating tube (lysing

beads with 750 μL lysing solution) provided in the PowerSoil DNA Isolation Kit (Mo-Bio Laboratories Inc., Carlsbad CA). Cell lysis by multidirectional beating was conducted in the FastPrep-24 homogenizer (MP Biomedicals LLC, Solon OH), following manufacturer recommendations of 30 seconds at 5.0 m/s. DNA was eluted in 50 μL solution C6, quantified using PicoGreen dsDNA assay (Invitrogen Life Technologies, Grand Island, NY), diluted to equimolar aliquots, and stored at -20oC until sequencing.

Bacterial DNA was analyzed at the Genomic Sequencing and Analysis Facility (GSAF) at the University of Texas at Austin (Austin, TX, USA) for Illumina® paired-end (2×250) sequencing on the MiSeq platform. For bacteria, first-round PCR was used to amplify the V4/V5 regions of the 16S rRNA gene using the primers 515F (5'-GTGYCAGCMGCCGCGGTA-3') (Baker et al., 2003) and 909R (5'-CCCCGYCAATTCMTTTRAGT-3') (Wang & Qian, 2009). Primers included appropriate Illumina adapters with reverse primers also having an error correcting 12-bp barcode unique to each sample to permit multiplexing of samples PCR amplification was performed using Qiagen Taq polymerase (Qiagen Corporation, Valencia, CA). After the PCR amplification, samples were prepared for their Illumina® sequencing run. This first round of PCR amplification was run in triplicate for each sample, pooled, and then cleaned using AMPure beads (New England Biolabs, Ipswich, MA). Second-round PCR amplification was performed with different primers that added sample-specific barcodes. Both rounds of PCR amplification (a total of 30 cycles) used Taq polymerase NEB Q5 (New England Biolabs, Ipswich, MA). The final PCR products for each sample after both rounds of amplification were then size-purified by removing amplicons less than 300 bp in length using AMPure beads (New England Biolabs, Ipswich, MA) and quantified using PicoGreen (Life Technologies, Carlsbad, CA). Samples were then normalized by amplicon mass and pooled for the Illumina® run. In addition, a random subset of samples were assessed on an Agilent BioAnalyzer (Agilent Technologies, Santa Clara, CA) to ensure correct amplicon size.

Bacterial sequences were processed and analyzed using the Quantitative Insights Into Microbial Ecology (QIIME) pipeline (Caporaso et al., 2010). Paired-end reads were stitched using FLASH with default parameters. Sequences were quality-filtered (-q 19), and chimeras were removed via QIIME and USEARCH. High-quality sequences were clustered into operational taxonomic units (OTUs) at 97% similarity using QIIME's USEARCH-based open-reference OTU clustering workflow (pick_open_reference_otus.py). Global singleton OTUs were removed, and OTU proportions were standardized to the total number of high-quality reads. Taxonomy was assigned using the Ribosomal Database Project classifier (Wang et al 2007) with the reference database Greengenes13_8 16s rRNA (Mcdonald D et al 2012) for bacteria. Because reads occurring in sample blanks were presumed to arise from background sequences on the filter, rather than cross-contamination during processing and extraction, reads obtained for each of the negative control OTUs were subtracted from the corresponding OTUs in each respective sample type.

2.4. SEQUENCE PROCESSING AND STATISTICAL ANALYSIS

Sequences were processed in the QIIME environment (Caporaso et al., 2010). All analyses were performed in the R environment (www.r-project.org). Pair-wise dissimilarities between communities were calculated using weighted UniFrac (Lozupone and Knight, 2005). Microbial community analysis of variance (implemented as ADONIS) and dispersion (implemented as betadisper) as well as mantel tests employed the Vegan package in R (Oksanen, 2016).

3. Results

3.1. Bacterial community within-sample diversity

After quality filtering, a total of 2,913,611 high-quality bacterial sequences were clustered into 40,563 operational taxonomic units (OTUs). Median sequences generated per sample were 81,986, 92,134, and 146,620 for vacuum, swab, and cut samples, respectively. All samples were rarefied to 51,290 reads, yielding a total of 36,265 distinct OTUs.

In total, two archaeal and 39 bacterial phyla were detected. Bacteria contributed an average of 99.9% of the sequences across all techniques, while archaea contributed 0.057 %, and < 0.002 % of the sequences were unidentified. Proteobacteria was the most abundant phylum recovered by all three techniques, representing from 40% of sequences in the case of vacuum samples to 45% in the case of swab samples. The next three most abundant phyla were Actinobacteria, Firmicutes, and Bacteroidetes representing ~26%, ~12%, and ~10% of reads, respectively (Fig 1).

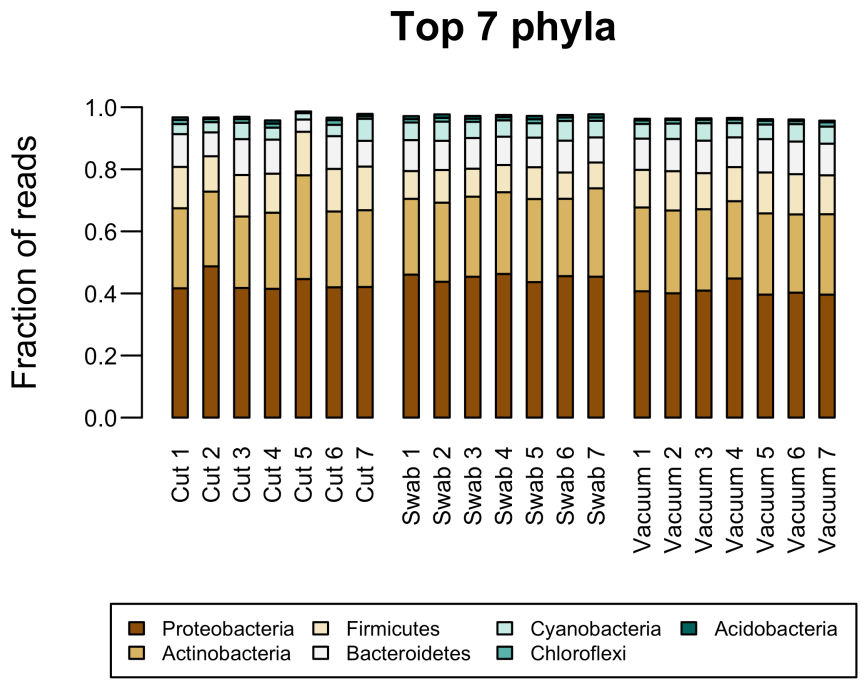

Figure 1: Relative abundance of bacterial phyla per sample

The ten most abundant OTUs recovered by each sample type are shown in Table 1. The top 5 OTUs are highlighted, and all OTUs are ranked to help visualize the degree of agreement between techniques. The ten most abundant OTUs in the cut samples represented a higher percentage (0.19%) than those in the swab (0.14%) and vacuum (0.15%) samples, respectively. As shown in Table 1, 6 of the top 10 OTUs are shared by all three sample types, and 8 are shared by at least 2 of the 3 sample types. The most abundant OTU, classified in the NCBI database as *Psychrobacter phenylpyruvicus*, is the same for all three sample types. This species has been isolated from human clinical material (Snell et al., 1972) and did not appear in the blank or negative control samples.

Greengenes genus*	Closest 16s NCBI isolate	Accesion	Similarity to isolate (%)	Rank OTUs Swabs	Rank OTUs Vacuum	Rank OTUs Cuts
Psychrobacter	Psychrobacter phenylpyruvicus	EU915471.1	99	1	1	1
Sphingomonas	Sphingomonas abaci	JX067906.1	100	2	3	2
Order Streptophyta	Tetrastigma hemsleyanum chloroplast	KT033563.1	100	3	2	6
Sphingomonas	Sphingomonas melonis	KT347493.1	100	4	4	5
Sphingomonas	Sphingomonas sp. NCCP-1295	LC065322.1	100	5	6	8
Family Oxalobacteraceae	Massilia aurea	KP318043.1	100	6	8	--
Sphingomonas	Sphingomonas phyllosphaerae	NR_029111.1	100	7	--	--
Skermanella	Skermanella aerolata	JX841089.1	100	8	--	--
Sphingomonas	Sphingomonas yunnanensis	JQ660311.1	100	9	--	--
Curtobacterium	Curtobacterium flaccumfaciens	KP296213.1	100	10	9	9
Micrococcus	Micrococcus luteus	KT448594.1	100	--	5	--
Corynebacterium	Corynebacterium tuberculostearicum	LN867524.1	100	--	7	--
Order Streptophyta	Merostachys Greco chloroplast	KT373815.1	100	--	10	7
Methylobacterium	Methylobacterium adhaesivum	AB698722.1	100	--	--	3
Corynebacterium	Corynebacterium ureicelerivorans	CP009215.1	100	--	--	4
Agrobacterium	Rhizobium sp. E51	KR703542.1	100	--	--	10

Table 1: Top 10 OTUs. The closest NCBI isolate, along with its accession number, and the similarity to that isolate are provided. The greengenes genus is also listed for comparison (DeSantis et al., 2006). In some cases, the OTUs could not be identified to the genus level. In those instances, the lowest taxa level obtained with greengenes is provided.

To further elucidate alpha diversity patterns, the cumulative distributions of reads per OTU were plotted for each of the three sample types (Fig 2). OTUs for each sample type were ranked in order of mean read abundance, and standard errors were computed for each OTU (n=7 for each sample type). Aggregated vacuum samples showed the greatest richness, followed by swab and cut samples (24,279, 18,363, and 14,518 OTUs, respectively, in the rarefied dataset). This high degree of richness reflects the large amount of biomass captured in HVAC filter dust. All sample types showed qualitatively similar long-tailed distributions, as 50% of the rarefied reads were captured by 0.64%, 0.68%, and 0.70% of the OTUs for cut, swab, and vacuum samples, respectively.

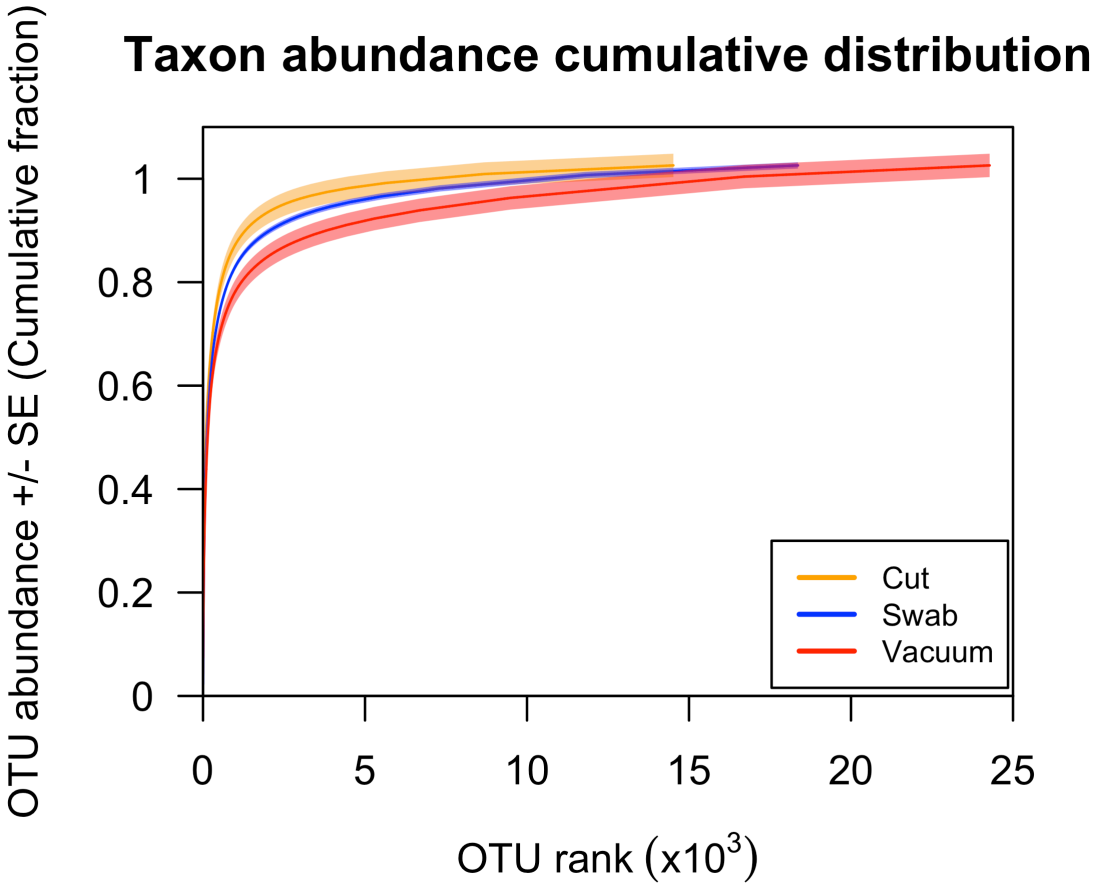

Figure 2: Cumulative distribution of reads per abundance-ranked OTU. Mean abundances within sample types (vacuums, swabs, and cuts) were computed for each OTU in the rarefied OTU table. OTUs were then ranked by mean abundance. Cumulative sums of these means were then normalized on a 0-1 scale. Center lines represents mean abundances. Shaded regions are standard errors based on 7 samples per sample type.

3.2. BACTERIAL SOURCE ENVIRONMENT ATTRIBUTION

Bacterial community composition was further broken down into taxonomic groupings indicative of potential source environments. The 11 families previously identified as human indicators (Corynebacteriaceae, Staphylococcaceae, Streptococcaceae, Lactobacillaceae, Propionibacteriaceae, Peptostreptococcaceae, Bifidobacteriaceae, Micrococcaceae, Dietziaceae, Aerococcaceae, Tissierellaceae) (Meadow et al., 2015) together accounted for 9, 5, and 10% of the reads for vacuum,

swab, and cut samples, respectively (Fig 3). Skin indicator genera (Barberan et al. 2015) contributed a greater proportion of reads than any other indicator group assessed. While vacuum and cut samples showed similar contributions from skin-associated genera, 7.5 and 6%, respectively, the same genera accounted for only ~3% reads in swab samples (Fig 4). Stool-associated genera (Barberan et al. 2015) in vacuum and cut samples contributed an average of 4% and 5% of the reads, while <3% in swab samples (Fig 5). Bacteria potentially sourced from soil, marine environments, and insects (Barberan et al. 2015) were also detectable, though in lower proportions.

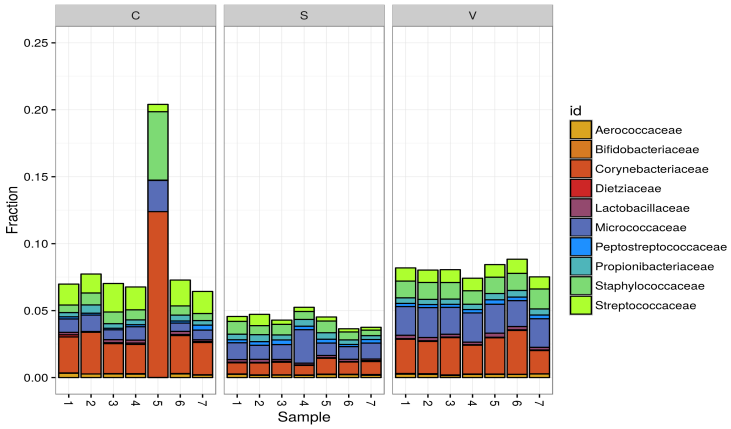

Figure 3: Human indicator bacterial families used in Meadows et al., 2015. The ordinate represents relative abundance, or "Fraction" of reads for a given sample. Note that Tissierellaceae is not reported in our open reference OTU picking.

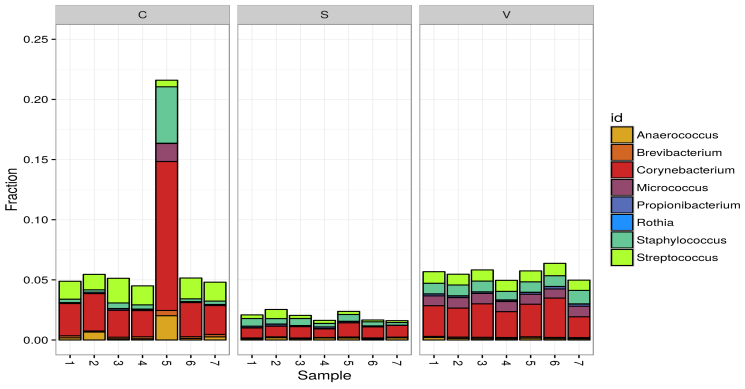

Figure 4: Skin indicator bacterial genera used in Barberan et al., 2015.

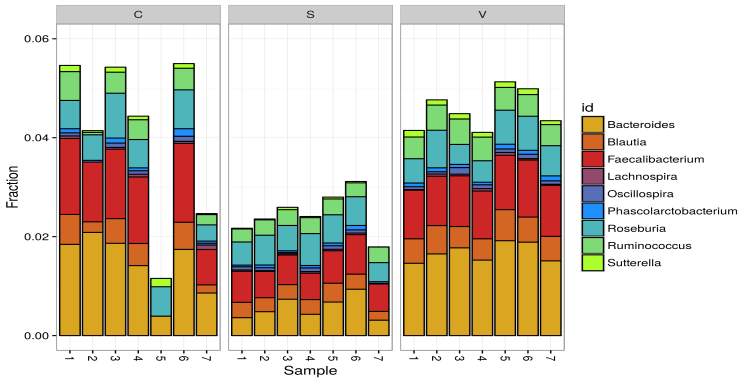

Figure 5: Human stool indicator genera used in Barberan et al., 2015.

3.3. BACTERIAL COMMUNITY BETWEEN-SAMPLE DIVERSITY

Distances between microbial assemblages were visualized with principal coordinate analysis plots (PCoA) and kernel density distributions (Fig 6). Communities were a priori grouped by sample type and analyzed with permutational analysis of dispersion and permutational analysis of variance (betadisper and ADONIS, respectively, vegan package, R). As can be observed in Fig 6, Weighted UniFrac distances (Lozupone and Knight, 2005) suggest that the bacterial assemblages recovered clustered by sample method, and that assemblages recovered by swabbing and vacuuming were less variable than those by cutting (global p=0.02, betadisper). Sample type explained about half of the variation between bacterial communities (ADONIS R2=0.48, p=0.001 on 999 permutations). It should be noted that this analysis may be affected by cut sample 5, which appears as an outlier in the source attribution plots above and in the following PCoA plot. This sample somewhat inflates the heteroscedasticity across sample types cited above.

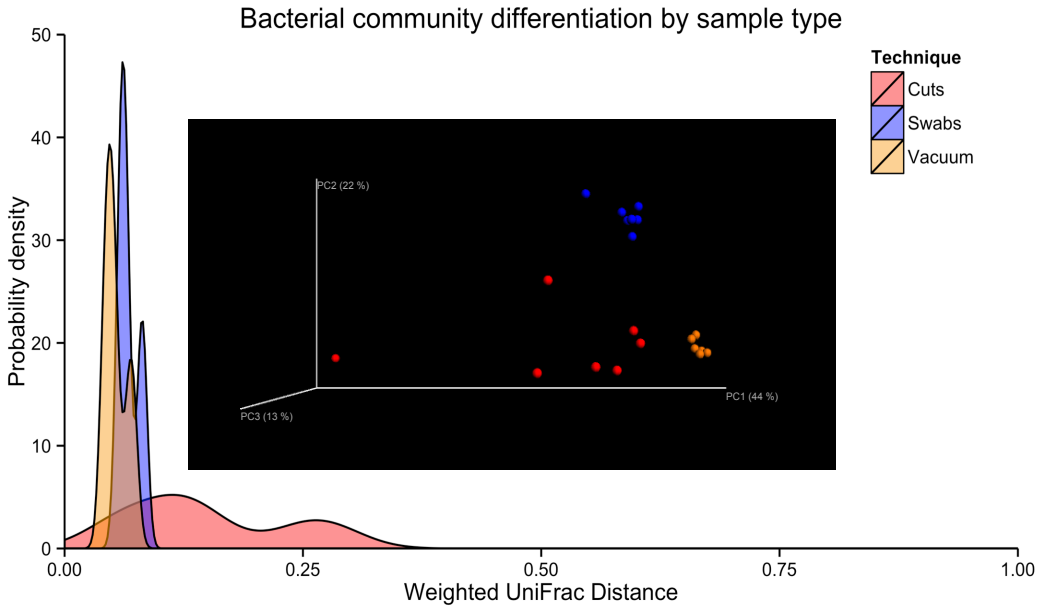

Figure 6: Community differentiation by sample type. Kernel density distributions were computed in R for weighted UniFrac distances between samples of the same sample type. Default variance and smoothing options were used. Inset is a PCoA plot colored by sample type.

The core microbiome concept was used to further assess the repeatability of each of the three techniques. Defined here as the OTUs shared by all samples of a given sample type, (e.g., the OTUs that appear in all seven vacuum samples), the core microbiome describes shared OTU membership. The core microbiome of vacuum samples represented 6.1 % of the total OTUs recovered in all vacuum samples. These OTUs accounted for 80% of the vacuum sample sequences obtained in the rarefied dataset. Similarly, the core microbiome for swab samples represented 6.3% of the total swab OTUs and 75% of the reads. Cut samples, however, recovered a smaller core comprising 1.6% of the total cut OTUs and 44% of the sequences. Thus, vacuums and swabs shared higher percentages of OTUs and sequences between all their respective samples than did cuts. It is also interesting to note that out of 36,265 bacterial OTUs in our whole data set, only 6,203 OTUs were common to all three sample types. Together, these numbers suggest that community structure is much more repeatable than

community membership, which is in part attributable to the long-tail OTU rank distributions depicted previously.

To understand how well the communities captured by each individual sample type reflected the global HVAC community approximated by the entire dataset, simplified communities comprising only core OTUs were constructed for each sample type. Weighted UniFrac distance matrices based on these three core communities were then correlated with the global distance matrix based on all OTUs using a permutational Mantel test. Results indicated that all three core bacterial matrices were highly correlated with the overall distance matrix, although vacuum- and swab-derived cores reflected patterns in the overall dataset (Mantel r=0.997 and 0.996, respectively; p=0.001 based on 999 permutations) even more strongly than did cut-derived core bacteria (Mantel r=0.95, p=0.001). Thus, core bacterial OTUs drove patterns in community structure, while sporadically present OTUs contributed little to community structure.

4. Discussion

Assessing the home microbiome precisely and accurately is essential in order to understand the potential interactions between the home microbiome and human health (Dannemiller et al., 2014, 2015, 2016; Valkonen et al., 2015). It is also key for investigating indoor microbial ecology (Adams et al., 2013; Barberán et al., 2015b; Kembel et al., 2012) and for potential applications such as microbiome forensics (Fierer et al., 2010; Lax et al., 2015a; Meadow et al., 2014a) and developing healthier buildings (Green, 2014; Kembel et al., 2014; Lax et al., 2015b; Meadow et al., 2014a). Since HVAC filters hold promise as long-term, airborne microbial samplers (Noris et al., 2011), the aim of this study was to examine various methods of dust removal. Specifically, three methods of removing dust from home HVAC filters were tested for their ability to produce repeatable, representative bacterial communities.

All sampling techniques revealed taxa typical of the home microbiome. The dominant bacterial phyla captured by the analyzed HVAC filter are characteristic of home microbiota described previously (Barberán et al., 2015b; Lax et al., 2014). These taxa are generally associated with humans, supporting the claim that humans are important sources of indoor airborne bacteria (Hospodsky et al., 2012, 2015; Meadow et al., 2015; Täubel et al., 2009).

All three sampling techniques also produced qualitatively similar long-tailed OTU distributions. Previous studies have found that more abundant OTUs tend to appear more frequently across samples than less abundant OTUs, which is characteristic of a random sampling process (Adams et al., 2015; Bowen et al., 2012; Fahlgren et al., 2010; Flores et al., 2014; Yamamoto et al., 2014). These long-tailed OTU distributions help explain the high bacterial richness observed here, and the fact that a relatively low portion of OTUs and high portion of reads were shared within sample types. Thus, investigating the occurrence of rare taxa across space may be difficult, especially if sequencing depth is

limited (Knight et al., 2012). Novel statistical methods may be needed to understand the potential importance of rare taxa for bacterial community dynamics (Shade et al., 2014). Given the higher repeatability of weighted measures of community structure, our findings suggest that environmental differences in community structure may be compared with greater confidence than differences in community membership.

Furthermore, our results show that vacuum and swab samples of HVAC filter dust were more repeatable than cut samples in terms of both community structure and membership. Community structure was more consistent across vacuum and swab samples as indicated by smaller weighted UniFrac distances and larger abundance-based core communities. Larger membership-based core OTU communities for these two samples types compared to cut samples showed that their community membership was also more consistent. Also, given that vacuum and swab core communities were more closely correlated than the vacuum core community to the global communities sampled by all three sampling methods, vacuum and swab HVAC filter samples appear to produce more representative HVAC samples. We speculate that this trend may extend beyond HVAC filter dust samples to other home environmental dust samples as well.

Practical advantages of HVAC filter sampling for indoor airborne microbiome analysis include the relative ease of installation and lack of intrusiveness during long-term sampling periods. Developing sampling protocols to remove dust from the filters depends not only on microbiological considerations but also on the relative cost, labor, and hardware of each method. When considering various passive samplers for airborne microbiome characterization in homes, one study found that microbial communities were little affected by sampler type when compared to effects from different environments, and thus ease of sample collection and economics were likely to be primary drivers of sampling protocol selection (Adams et al., 2015). In our study, cut sampling stands out as being considerably more labor intensive than vacuum and swab sampling, since cutting HVAC filters is physically difficult, and the following elution and filtration steps require

additional time. Cut sampling is thus not recommended based on both bacterial community repeatability and practical considerations.

A limitation of our study is that it does not distinguish technical sampling variability from environmental variability. A permutational analysis of variance-type method showed that about half of the variation across communities was attributable to sample type. A previous study that distinguished between these sources of variability in home floor dust bacterial communities collected by vacuuming, found that technical replicates were highly repeatable, while dust samples collected from adjacent 1 m^2 areas were more variable, though still highly concordant (Fujimura et al., 2012). Our study was designed to minimize environmental variability by compositing dust from five randomly chosen filter locations on a single filter. However, we surmise that variability between technical replicates and sample types would be relatively insignificant compared to the variability encountered across different filters from different homes. This will be tested by follow up analysis of vacuum and swab samples taken from a set of different homes.

5. Conclusions

Our findings suggest that vacuum and swab samples produced more repeatable and representative bacterial communities than did elution. Furthermore, given the reduced labor and cost of vacuum and swab methods, and the additional advantage that these two methods may also be applied to sampling dust from other home surfaces, vacuum and swab sampling of HVAC filter dust are found to be superior to elution. Future work will expand the scope of these findings by examining the relative importance of technical variability associated with these sampling methods compared to the environmental variability associated with sampling HVAC filters from different homes. Analysis will also be expanded to assess method and environmental variability for other measures as well, including fungal communities, and quantitative measures of total bacteria and fungi.

Appendix: R scripts written for plotting and analysis

```
############

#Plot Top 7 bacterial phyla

############

library(RColorBrewer)

p <- read.csv("~/Box Sync/Experimental_Techniques/Bacterial
Analysis/percentagesBactArchUnidentified_table_mc51290_sorted_L2/Plot_top_7_phyl
a.csv")
head(p)
str(p)

Rep <- c(1:7)
Type <- c(rep(c("Cut", "Swab", "Vacuum"), each=7))
names <- paste(Type, Rep, sep=" ")

#Stacked barplot
space.vec <- c(0,1,1,1,1,1,1,2,1,1,1,1,1,1,2,1,1,1,1,1,1)
cols <- (brewer.pal(7, "BrBG"))

 w = 1.2
 r = 1.7
png("~/Box Sync/Experimental_Techniques/Bacterial
Analysis/percentagesBactArchUnidentified_table_mc51290_sorted_L2/Barplot_top_7_p
hyla.png", height=1920, width=1920*w, res=288*r)

par(xpd=TRUE, mar=c(8,5,4,1))
barplot(as.matrix(p[,-1]), main="Top 7 phyla", names.arg=names, cex.names=0.7,
cex.axis=0.7, ylab="Fraction of reads", ylim=c(0,1), las=2, col=cols, space=space.vec,
cex.lab=1, , mgp=c(2.5,.6,0.1))
legend(x=-4, y=-.6, legend=p$Phylum, cex=.6, col=cols, fill=cols, horiz=F, ncol=4)

    dev.off()


##############

#Create cumulative distribution plots for reads vs OTUs by sample type

##############

#Bacteria
#Create cumulative distribution plots for reads vs OTUs by sample type
```

```
#use the OTU ID table rarefied to 51290, instead of the OTU table with taxa assigned because this only gives ~1200 OTUs (due to clustering of OTUs into larger "unknown" categories?)
# a=data<- read.delim("c:/Users/jm44736/Box Sync/Microbiome-HUD/Experimental_Techniques/Bacterial Analysis/otu_table_mc2_w_tax_cleaned_final_wtax_51290_readsotu.txt")
a <- read.delim("~/Box Sync/Experimental_Techniques/Bacterial Analysis/otu_table_mc2_w_tax_cleaned_final_wtax_51290_readsotu.txt")
head(a)
str(a)
#qqnorm(log(a[1, grepl("Cut.\\d", colnames(a))]+1))

#Means of OTU abundances by sample type

a$cut.av <- rowMeans(a[, grepl("Cut.\\d", colnames(a))])
a$swab.av <- rowMeans(a[, grepl("Swab.\\d", colnames(a))])
a$vac.av <- rowMeans(a[, grepl("Vacuum.\\d", colnames(a))])
min(a$cut.av)
max(a$cut.av)
sum(a$cut.av)

#Standard error of OTU abundances by sample type

se <- function(x) {
        sd(x)/sqrt((length(x)-1))
}

a$cut.se <- apply(a[, grepl("Cut.\\d", colnames(a))], 1, se)
a$swab.se <- apply(a[, grepl("Swab.\\d", colnames(a))], 1, se)
a$vac.se <- apply(a[, grepl("Vacuum.\\d", colnames(a))], 1, se)
range(a$cut.se)
head(a)

t.val <- 1 #2.447 # two side alpha =0.05, t value for 6 df

#Create new data frame with OTUs ranked by abundance. OTU IDs will be replaced by their rank, i.e., 1, 2, 3,... Thus, OTU #1 for cuts may no longer be the same as OTU #1 for swabs, as the most abundant OTU may be different across sample types

OTU.ranked <- data.frame(
        rank = seq(from=1, to=dim(a)[1], by=1),
        cut = a$cut.av[order(a$cut.av, decreasing=T)],
        swab = a$swab.av[order(a$swab.av, decreasing=T)],
        vac = a$vac.av[order(a$vac.av, decreasing=T)])
OTU.ranked$cut.cum = cumsum(OTU.ranked$cut)
```

```
OTU.ranked$swab.cum =  cumsum(OTU.ranked$swab)
OTU.ranked$vac.cum =  cumsum(OTU.ranked$vac)
OTU.ranked$cut.cum.up = OTU.ranked$cut.cum +
t.val*(cumsum(a$cut.se[order(a$cut.av, decreasing=T)]^2))^(1/2)
OTU.ranked$cut.cum.lo = OTU.ranked$cut.cum -
t.val*(cumsum(a$cut.se[order(a$cut.av, decreasing=T)]^2))^(1/2)
OTU.ranked$swab.cum.up = OTU.ranked$swab.cum +
t.val*(cumsum(a$swab.se[order(a$swab.av, decreasing=T)]^2))^(1/2)
OTU.ranked$swab.cum.lo = OTU.ranked$swab.cum -
t.val*(cumsum(a$swab.se[order(a$swab.av, decreasing=T)]^2))^(1/2)
OTU.ranked$vac.cum.up = OTU.ranked$vac.cum +
t.val*(cumsum(a$cut.se[order(a$cut.av, decreasing=T)]^2))^(1/2)
OTU.ranked$vac.cum.lo = OTU.ranked$vac.cum -
t.val*(cumsum(a$cut.se[order(a$cut.av, decreasing=T)]^2))^(1/2)
        head(OTU.ranked)
        dim(OTU.ranked)


#Compute the minimum number of OTUs accounting for 50% of all reads

half.reads <- 0.5*51290
cut50th <- min(OTU.ranked[OTU.ranked$cut.cum>=half.reads, "rank"]); cut50th
swab50th <- min(OTU.ranked[OTU.ranked$swab.cum>=half.reads, "rank"]); swab50th
vac50th <- min(OTU.ranked[OTU.ranked$vac.cum>=half.reads, "rank"]); vac50th
vacTop1per <- OTU.ranked$vac.cum[240]; vacTop1per/51290

plot.cut <- unique(OTU.ranked$cut.cum)
plot.swab <- unique(OTU.ranked$swab.cum)
plot.vac <- unique(OTU.ranked$vac.cum)

OTU.ranked$rank.frac.cut <- OTU.ranked$rank/length(plot.cut)
OTU.ranked$rank.frac.swab <- OTU.ranked$rank/length(plot.swab)
OTU.ranked$rank.frac.vac <- OTU.ranked$rank/length(plot.vac)

per.cut50th <- cut50th/length(plot.cut); per.cut50th
per.swab50th <- swab50th/length(plot.swab); per.swab50th
per.vac50th <- vac50th/length(plot.vac); per.vac50th


#Plot the cumulative distribution, reads normalized (by changing the axis dimensions)
and OTUs not normalized

 w = 1.2
 r = 1.7
```

```
png("~/Box Sync/Experimental_Techniques/Bacterial Analysis/Reads per OTU
distribution/Cumulative_frac_dist_reads_OTU_bacteria_w_SE.png", height=1920,
width=1920*w, res=288*r)

par(mar=c(5,5,3,2))
plot(OTU.ranked$cut.cum~OTU.ranked$rank, type="n", xlim=c(0,25000),
ylim=c(0,55000), xaxs="i", xaxt="n", yaxs="i", yaxt="n", ylab=expression("OTU
abundance +/- SE (Cumulative fraction)"), xlab=expression("OTU rank" ~ (x10^{3})),
main="Taxon abundance cumulative distribution")
lines(plot.cut~OTU.ranked$rank[1:length(plot.cut)], lty=1, lwd=1, col="orange")
lines(plot.swab~OTU.ranked$rank[1:length(plot.swab)], lty= 1, lwd =1, col="blue")
lines(plot.vac~OTU.ranked$rank[1:length(plot.vac)], lty=1, lwd=1, col="red")

axis(1, at = seq(0, 25000, by=5000), labels = seq(0, 25, by = 5))
axis(2, at = seq(0, 50000, by=10000), labels = seq(0, 1, by = .2), las=2)

polygon(c(OTU.ranked$rank[1:length(plot.cut)], rev(OTU.ranked
                        $rank[1:length(plot.cut)])),
c(OTU.ranked$cut.cum.up[1:length(plot.cut)],
        rev(OTU.ranked$cut.cum.lo[1:length(plot.cut)])),
    col = rgb(t(col2rgb("orange"))/255, alpha=0.5), border = NA)
polygon(c(OTU.ranked$rank[1:length(plot.swab)], rev(OTU.ranked
                              $rank[1:length(plot.swab)])),
c(OTU.ranked$swab.cum.up[1:length(plot.swab)],
        rev(OTU.ranked$swab.cum.lo[1:length(plot.swab)])),
    col = rgb(t(col2rgb("blue"))/255, alpha=0.5), border = NA)
polygon(c(OTU.ranked$rank[1:length(plot.vac)], rev(OTU.ranked
                        $rank[1:length(plot.vac)])),
c(OTU.ranked$vac.cum.up[1:length(plot.vac)],
        rev(OTU.ranked$vac.cum.lo[1:length(plot.vac)])),
    col = rgb(t(col2rgb("red"))/255, alpha=0.5), border = NA)

legend(17550, 15000, c("Cut", "Swab", "Vacuum"), lty=c(1,1,1), col=c("orange", "blue",
"red"), lwd=c(2,2,2), cex=.75)


     dev.off()


                                    ##############

                            #Weighted UniFrac distance density plots

                                    ##############

# install.packages("vegan")
library(ggplot2)
```

```
a <- read.table("~/Box Sync/Experimental_Techniques/Bacterial Analysis/Beta diversity
analysis/Distance matrices/weighted_unifrac_dm.txt", header=T, sep="")
head(a)
str(a)

#Create dataframes for all "cut-to-cut" distances called "cuts"; same for swabs and vacs
cuts <- data.frame()
cuts <- a[grepl("Cut.\\d", names(a)),grepl("Cut.\\d", names(a))]
swabs <- data.frame()
swabs <- a[grepl("Swab.\\d", names(a)),grepl("Swab.\\d", names(a))]
vacs <- data.frame()
vacs <- a[grepl("Vacuum.\\d", names(a)),grepl("Vacuum.\\d", names(a))]

#Restructure these dataframes so that they contain only 1 column each; remove duplicate
values (each value occurs twice) and zeroes
cutsvec <- as.vector(as.matrix(cuts))
cutsvec <- cutsvec[cutsvec != 0]
cutsvec <- unique(cutsvec)
cutsvec.se <- sd(cutsvec)/sqrt(length(cutsvec)-1)
unifrac.min.detection.diff.c <- mean(cutsvec) + 2*cutsvec.se
cutsdens <- density(cutsvec)

swabsvec <- as.vector(as.matrix(swabs))
swabsvec <- swabsvec[swabsvec != 0]
swabsvec <- unique(swabsvec)
swabsvec.se <- sd(swabsvec)/sqrt(length(swabsvec)-1)
unifrac.min.detection.diff.s <- mean(swabsvec) + 2*swabsvec.se
swabsdens <- density(swabsvec)

vacsvec <- as.vector(as.matrix(vacs))
vacsvec <- vacsvec[vacsvec != 0]
vacsvec <- unique(vacsvec)
vacsvec.se <- sd(vacsvec)/sqrt(length(vacsvec)-1)
unifrac.min.detection.diff.v <- mean(vacsvec) + 2*vacsvec.se
vacsdens <- density(vacsvec)

within <- data.frame(dist = c(vacsvec, swabsvec, cutsvec), lines=rep(c("Vacuum",
"Swabs", "Cuts"), each=length(vacsvec)))
within
fill=c("red", "blue", "orange")
ggplot(within, aes(x=dist, fill=lines)) +
        geom_density(alpha=0.5) +
        scale_fill_manual(values = c("red", "blue", "orange"),

 name="Technique") +
```

```
        scale_x_continuous(limits = c(0, 1), expand = c(0, 0)) +
        scale_y_continuous(limits = c(0, 50), expand = c(0, 0)) +
        ggtitle("Bacterial community differentiation by sample type") +
        ylab("Probability density") + xlab("Weighted UniFrac Distance") +
        theme_classic() +
    theme(legend.position=c(0.9,0.85))



                                    #############

        #BETA DIVERSITY STATS: Weighted UniFrac distance, ADONIS, Mantel

                                    #############

# install.packages("vegan")
library(vegan)
library(ggplot2)

a <- read.table("~/Box Sync/Experimental_Techniques/Bacterial Analysis/Beta diversity
analysis/Distance matrices/weighted_unifrac_dm.txt", header=T, sep="")
head(a)
str(a)

#Create distance matrix excluding controls for all distances, grouping vector, and run
ANOSIM

dist <- data.frame()
tomatch <- c("Cut.\\d", "Swab.\\d", "Vacuum.\\d")
dist <- a[grepl(paste(tomatch, collapse="|"), names(a)), grepl(paste(tomatch,
collapse="|"), names(a))]

#order dist matrix alphabetically
dist <- dist[order(names(dist)), order(names(dist))]

#create group vector as factor for ANOVA-type tests
group <- as.factor(ifelse(grepl("Cut.\\d", row.names(dist))==T, "Cut",
ifelse(grepl("Swab.\\d", row.names(dist))==T, "Swab", "Vacuum")))

#The distance matrix must be converted to class 'dist', which simply makes it a lower
diagonal matrix without repeat or zero values
dist <- as.dist(dist)

#Run betadisp to check for homogeneity of variance by sample type
mod <- betadisper(dist, group) #Globally, it appears as though there may be a difference
in variances
boxplot(mod)
```

```
permutest(mod)
TukeyHSD(mod) #pairwise, however, there's not much evidence to support a difference in variances by sample type
#Thus, can run ADONIS, but part of the variance attributaible to the ADONIS model may be due to the difference in spread shown here.

#Run adonis as more robust alternative to ANOSIM;
str(dist)
dist
adonis(dist ~ group, perm=999)

#anosim(dist, group, permutations = 1000)

#Run mantel test using pearson's correlation between core microbiome distance matrix and all-samples distance matrix

all <- read.table("~/Box Sync/Experimental_Techniques/Bacterial Analysis/Beta diversity analysis/Distance matrices/weighted_unifrac_dm.txt", sep="", header=T)
all <- as.dist(all[order(names(all)), order(names(all))])

cut.core <- read.table("~/Box Sync/Experimental_Techniques/Bacterial Analysis/otu_table_core_cut/weighted_unifrac_core_table_100.txt", sep="", header=T)
tomatch <- c("Cut.\\d", "Swab.\\d", "Vacuum.\\d")
cut.core <- cut.core[grepl(paste(tomatch, collapse="|"), names(cut.core)), grepl(paste(tomatch, collapse="|"), names(cut.core))]
cut.core <- as.dist(cut.core[order(names(cut.core)), order(names(cut.core))])

swab.core <- read.table("~/Box Sync/Experimental_Techniques/Bacterial Analysis/otu_table_core_swab/weighted_unifrac_core_table_100.txt", sep="", header=T)
swab.core <- swab.core[grepl(paste(tomatch, collapse="|"), names(swab.core)), grepl(paste(tomatch, collapse="|"), names(swab.core))]
swab.core <- as.dist(swab.core[order(names(swab.core)), order(names(swab.core))])

vac.core <- read.table("~/Box Sync/Experimental_Techniques/Bacterial Analysis/otu_table_core_vacuum/weighted_unifrac_core_table_100.txt", sep="", header=T)
vac.core <- vac.core[grepl(paste(tomatch, collapse="|"), names(vac.core)), grepl(paste(tomatch, collapse="|"), names(vac.core))]
vac.core <- as.dist(vac.core[order(names(vac.core)), order(names(vac.core))])

#Run Mantel
mantel(all, cut.core, method="pearson", permutations=999)
mantel(all, swab.core, method="pearson", permutations=999)
mantel(all, vac.core, method="pearson", permutations=999)

plot(cut.core, all)
```

References

Adams, R.I., Miletto, M., Taylor, J.W., and Bruns, T.D. (2013). Dispersal in microbes: fungi in indoor air are dominated by outdoor air and show dispersal limitation at short distances. ISME J. 7, 1262–1273.

Adams, R.I., Miletto, M., Lindow, S.E., Taylor, J.W., and Bruns, T.D. (2014). Airborne Bacterial Communities in Residences: Similarities and Differences with Fungi. PLOS ONE 9, e91283.

Adams, R.I., Tian, Y., Taylor, J.W., Bruns, T.D., Hyvärinen, A., and Täubel, M. (2015). Passive dust collectors for assessing airborne microbial material. Microbiome 3.

Adgate, J.L., Eberly, L.E., Stroebel, C., Pellizzari, E.D., and Sexton, K. (2004). Personal, Indoor, and Outdoor VOC Exposures in a Probability Sample of Children. J. Expo. Sci. Environ. Epidemiol. 14, S4–S13.

Andelman, J.B. (1985). Organic Micropollutants in Drinking Water and Health Inhalation exposure in the home to volatile organic contaminants of drinking water. Sci. Total Environ. 47, 443–460.

Andrew Hoisington, J.P.M. (2014). The Impact of Sampler Selection on Characterizing the Indoor Microbiome. Build. Environ.

Barberán, A., Ladau, J., Leff, J.W., Pollard, K.S., Menninger, H.L., Dunn, R.R., and Fierer, N. (2015a). Continental-scale distributions of dust-associated bacteria and fungi. Proc. Natl. Acad. Sci. 112, 5756–5761.

Barberán, A., Dunn, R.R., Reich, B.J., Pacifici, K., Laber, E.B., Menninger, H.L., Morton, J.M., Henley, J.B., Leff, J.W., Miller, S.L., et al. (2015b). The ecology of microscopic life in household dust. In Proc. R. Soc. B, (The Royal Society), p. 20151139.

Barnes, C.S., Allenbrand, R., Mohammed, M., Gard, L., Pacheco, F., Kennedy, K., Portnoy, J.M., and Ciaccio, C. (2015). Measurement of aeroallergens from furnace filters. Ann. Allergy. Asthma. Immunol. 114, 221–225.

Bartram, A.K., Lynch, M.D.J., Stearns, J.C., Moreno-Hagelsieb, G., and Neufeld, J.D. (2011). Generation of Multimillion-Sequence 16S rRNA Gene Libraries from Complex Microbial Communities by Assembling Paired-End Illumina Reads. Appl. Environ. Microbiol. 77, 3846–3852.

Berglund, B., Brunekreef, B., Knöppe, H., Lindvall, T., Maroni, M., Mølhave, L., and Skov, P. (1992). Effects of Indoor Air Pollution on Human Health. Indoor Air 2, 2–25.

Bernstein, J.A., Alexis, N., Bacchus, H., Bernstein, I.L., Fritz, P., Horner, E., Li, N., Mason, S., Nel, A., Oullette, J., et al. (2008). The health effects of nonindustrial indoor air pollution. J. Allergy Clin. Immunol. 121, 585–591.

Bowen, J.L., Morrison, H.G., Hobbie, J.E., and Sogin, M.L. (2012). Salt marsh sediment diversity: a test of the variability of the rare biosphere among environmental replicates. ISME J. 6, 2014–2023.

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Pena, A.G., Goodrich, J.K., Gordon, J.I., et al. (2010). QIIME allows analysis of high-throughput community sequencing data. Nat. Methods 7, 335–336.

Corsi, R.L., Kinney, K.A., and Levin, H. (2012). Microbiomes of built environments: 2011 symposium highlights and workgroup recommendations. Indoor Air 22, 171–172.

Dannemiller, K.C., Mendell, M.J., Macher, J.M., Kumagai, K., Bradman, A., Holland, N., Harley, K., Eskenazi, B., and Peccia, J. (2014). Next-generation DNA sequencing reveals that low fungal diversity in house dust is associated with childhood asthma development. Indoor Air 24, 236–247.

Dannemiller, K.C., Gent, J.F., Leaderer, B.P., and Peccia, J. (2015). Influence of housing characteristics on bacterial and fungal communities in homes of asthmatic children. Indoor Air.

Dannemiller, K.C., Gent, J.F., Leaderer, B.P., and Peccia, J. (2016). Indoor microbial communities: Influence on asthma severity in atopic and nonatopic children. J. Allergy Clin. Immunol.

DeSantis, T.Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E.L., Keller, K., Huber, T., Dalevi, D., Hu, P., and Andersen, G.L. (2006). Greengenes, a Chimera-Checked 16S rRNA Gene Database and Workbench Compatible with ARB. Appl. Environ. Microbiol. 72, 5069–5072.

Dunn, R.R., Fierer, N., Henley, J.B., Leff, J.W., and Menninger, H.L. (2013). Home Life: Factors Structuring the Bacterial Diversity Found within and between Homes. PLOS ONE 8, e64133.

Ege, M.J., Mayer, M., Normand, A.-C., Genuneit, J., Cookson, W.O.C.M., Braun-Fahrländer, C., Heederik, D., Piarroux, R., and von Mutius, E. (2011). Exposure to Environmental Microorganisms and Childhood Asthma. N. Engl. J. Med. 364, 701–709.

Emerson, J.B., Keady, P.B., Brewer, T.E., Clements, N., Morgan, E.E., Awerbuch, J., Miller, S.L., and Fierer, N. (2015). Impacts of Flood Damage on Airborne Bacteria and Fungi in Homes after the 2013 Colorado Front Range Flood. Environ. Sci. Technol. 49, 2675–2684.

Fahlgren, C., Bratbak, G., Sandaa, R.-A., Thyrhaug, R., and Zweifel, U.L. (2010). Diversity of airborne bacteria in samples collected using different devices for aerosol collection. Aerobiologia 27, 107–120.

Fierer, N., Lauber, C.L., Zhou, N., McDonald, D., Costello, E.K., and Knight, R. (2010). Forensic identification using skin bacterial communities. Proc. Natl. Acad. Sci. 107, 6477–6481.

Flores, G.E., Bates, S.T., Caporaso, J.G., Lauber, C.L., Leff, J.W., Knight, R., and Fierer, N. (2013). Diversity, distribution and sources of bacteria in residential kitchens. Environ. Microbiol. 15, 588–596.

Flores, G.E., Caporaso, J.G., Henley, J.B., Rideout, J.R., Domogala, D., Chase, J., Leff, J.W., Vázquez-Baeza, Y., Gonzalez, A., Knight, R., et al. (2014). Temporal variability is a personalized feature of the human microbiome. Genome Biol. 15.

Franck, U., Herbarth, O., Wehner, B., Wiedensohler, A., and Manjarrez, M. (2003). How do the indoor size distributions of airborne submicron and ultrafine particles in the absence of significant indoor sources depend on outdoor distributions? Indoor Air 13, 174–181.

Frankel, M., Timm, M., Hansen, E.W., and Madsen, A.M. (2012). Comparison of sampling methods for the assessment of indoor microbial exposure. Indoor Air 22, 405–414.

Fujimura, K.E., Rauch, M., Matsui, E., Iwai, S., Calatroni, A., Lynn, H., Mitchell, H., Johnson, C.C., Gern, J.E., Togias, A., et al. (2012). Development of a standardized approach for environmental microbiota investigations related to asthma development in children. J. Microbiol. Methods 91, 231–239.

Fujimura, K.E., Demoor, T., Rauch, M., Faruqi, A.A., Jang, S., Johnson, C.C., Boushey, H.A., Zoratti, E., Ownby, D., Lukacs, N.W., et al. (2014). House dust exposure mediates gut microbiome Lactobacillus enrichment and airway immune defense against allergens and virus infection. Proc. Natl. Acad. Sci. 111, 805–810.

Fung, F., and Hughson, W.G. (2003). Health Effects of Indoor Fungal Bioaerosol Exposure. Appl. Occup. Environ. Hyg. 18, 535–544.

Gordon, J., Gandhi, P., Shekhawat, G., Frazier, A., Hampton-Marcell, J., and Gilbert, J.A. (2015). A simple novel device for air sampling by electrokinetic capture. Microbiome 3, 79.

Green, J.L. (2014). Can bioinformed design promote healthy indoor ecosystems? Indoor Air 24, 113–115.

Hardin, B.D., Kelman, B.J., and Saxon, A. (2003). Adverse human health effects associated with molds in the indoor environment. J. Occup. Environ. Med. Coll. Occup. Environ. Med. 45, 470–478.

Hoisington, A., Maestre, J.P., Siegel, J.A., and Kinney, K.A. (2014). Exploring the microbiome of the built environment: A primer on four biological methods available to building professionals. HVACR Res. 20, 167–175.

Hoisington, A., Maestre, J.P., Kinney, K.A., and Siegel, J.A. (2015). Characterizing the bacterial communities in retail stores in the United States. Indoor Air n/a – n/a.

Hospodsky, D., Qian, J., Nazaroff, W.W., Yamamoto, N., Bibby, K., Rismani-Yazdi, H., and Peccia, J. (2012). Human Occupancy as a Source of Indoor Airborne Bacteria. PLOS ONE 7, e34867.

Hospodsky, D., Yamamoto, N., Nazaroff, W.W., Miller, D., Gorthala, S., and Peccia, J. (2015). Characterizing airborne fungal and bacterial concentrations and emission rates in six occupied children's classrooms. Indoor Air 25, 641–652.

Jari Oksanen, F.G.B. (2016). vegan Community Ecology Package Version 2.3-3. URL https://cran.r-project.org/web/packages/vegan/.

Jones, A.P. (1999). Indoor air quality and health. Atmos. Environ. 33, 4535–4564.

Kelley, S.T., and Gilbert, J.A. (2013). Studying the microbiology of the indoor environment. Genome Biol 14, 202.

Kembel, S.W., Jones, E., Kline, J., Northcutt, D., Stenson, J., Womack, A.M., Bohannan, B.J., Brown, G.Z., and Green, J.L. (2012). Architectural design influences the diversity and structure of the built environment microbiome. ISME J. 6, 1469–1479.

Kembel, S.W., Meadow, J.F., O'Connor, T.K., Mhuireach, G., Northcutt, D., Kline, J., Moriyama, M., Brown, G.Z., Bohannan, B.J.M., and Green, J.L. (2014). Architectural Design Drives the Biogeography of Indoor Bacterial Communities. PLoS ONE 9, e87093.

KLEPEIS, N.E., NELSON, W.C., OTT, W.R., ROBINSON, J.P., TSANG, A.M., SWITZER, P., BEHAR, J.V., HERN, S.C., and ENGELMANN, W.H. (2001). The National Human Activity Pattern Survey (NHAPS): a resource for assessing exposure to environmental pollutants. J. Expo. Sci. Environ. Epidemiol. 11, 231–252.

Knight, R., Jansson, J., Field, D., Fierer, N., Desai, N., Fuhrman, J.A., Hugenholtz, P., van der Lelie, D., Meyer, F., Stevens, R., et al. (2012). Unlocking the potential of metagenomics through replicated experimental design. Nat. Biotechnol. 30, 513–520.

Lax, S., Smith, D.P., Hampton-Marcell, J., Owens, S.M., Handley, K.M., Scott, N.M., Gibbons, S.M., Larsen, P., Shogan, B.D., Weiss, S., et al. (2014). Longitudinal analysis of microbial interaction between humans and the indoor environment. Science 345, 1048–1052.

Lax, S., Nagler, C.R., and Gilbert, J.A. (2015b). Our interface with the built environment: immunity and the indoor microbiota. Trends Immunol. 36, 121–123.

Lax, S., Hampton-Marcell, J.T., Gibbons, S.M., Colares, G.B., Smith, D., Eisen, J.A., and Gilbert, J.A. (2015a). Forensic analysis of the microbiome of phones and shoes. Microbiome 3, 21.

Lozupone, C., and Knight, R. (2005). UniFrac: a New Phylogenetic Method for Comparing Microbial Communities. Appl. Environ. Microbiol. 71, 8228–8235.

Luongo, J.C., Barberán, A., Hacker-Cary, R., Morgan, E.E., Miller, S.L., and Fierer, N. (2016). Microbial analyses of airborne dust collected from dormitory rooms predict the sex of occupants. Indoor Air n/a – n/a.

Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., Bemben, L.A., Berka, J., Braverman, M.S., Chen, Y.-J., Chen, Z., et al. (2005). Genome sequencing in microfabricated high-density picolitre reactors. Nature 437, 376–380.

Meadow, J.F., Altrichter, A.E., Kembel, S.W., Kline, J., Mhuireach, G., Moriyama, M., Northcutt, D., O'connor, T.K., Womack, A.M., Brown, G.Z., et al. (2014a). Indoor airborne bacterial communities are influenced by ventilation, occupancy, and outdoor air source. Indoor Air 24, 41–48.

Meadow, J.F., Altrichter, A.E., Kembel, S.W., Moriyama, M., O'Connor, T.K., Womack, A.M., Brown, G.Z., Green, J.L., and Bohannan, B.J. (2014b). Bacterial communities on classroom surfaces vary with human contact. Microbiome 2, 1.

Meadow, J.F., Altrichter, A.E., Bateman, A.C., Stenson, J., Brown, G.Z., Green, J.L., and Bohannan, B.J.M. (2015). Humans differ in their personal microbial cloud. PeerJ 3, e1258.

Noris, F., Siegel, J.A., and Kinney, K.A. (2011). Evaluation of HVAC filters as a sampling mechanism for indoor microbial communities. Atmos. Environ. 45, 338–346.

Peccia, J., Hospodsky, D., and Bibby, K. (2011). New Directions: A revolution in DNA sequencing now allows for the meaningful integration of biology with aerosol science. Atmos. Environ. 45, 1896–1897.

Quail, M.A., Smith, M., Coupland, P., Otto, T.D., Harris, S.R., Connor, T.R., Bertoni, A., Swerdlow, H.P., and Gu, Y. (2012). A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. BMC Genomics 13, 341.

Shade, A., Jones, S.E., Caporaso, J.G., Handelsman, J., Knight, R., Fierer, N., and Gilbert, J.A. (2014). Conditionally Rare Taxa Disproportionately Contribute to Temporal Changes in Microbial Diversity. mBio 5, e01371–14.

Shendure, J., and Ji, H. (2008). Next-generation DNA sequencing. Nat. Biotechnol. 26, 1135–1145.

Shendure, J., Porreca, G.J., Reppas, N.B., Lin, X., McCutcheon, J.P., Rosenbaum, A.M., Wang, M.D., Zhang, K., Mitra, R.D., and Church, G.M. (2005). Accurate multiplex polony sequencing of an evolved bacterial genome. Science 309, 1728–1732.

Snell, J.J.S., Hill, L.R., and Lapage, S.P. (1972). Identification and characterization of Moraxella phenylpyruvica. J. Clin. Pathol. 25, 959–965.

Spengler, J.D., and Sexton, K. (1983). Indoor air pollution: a public health perspective. Science 221, 9–17.

Stolwijk, J. a. J. (1992). Risk Assessment of Acute Health and Comfort Effects of Indoor Air Pollution. Ann. N. Y. Acad. Sci. 641, 56–62.

Täubel, M., Rintala, H., Pitkäranta, M., Paulin, L., Laitinen, S., Pekkanen, J., Hyvärinen, A., and Nevalainen, A. (2009). The occupant as a source of house dust bacteria. J. Allergy Clin. Immunol. 124, 834–840.e47.

Valkonen, M., Wouters, I.M., Täubel, M., Rintala, H., Lenters, V., Vasara, R., Genuneit, J., Braun-Fahrländer, C., Piarroux, R., Mutius, E. von, et al. (2015). Bacterial Exposures and Associations with Atopy and Asthma in Children. PLOS ONE 10, e0131594.

Wu, N., Herrmann, T., Paepke, O., Tickner, J., Hale, R., Harvey, E., La Guardia, M., McClean, M.D., and Webster, T.F. (2007). Human Exposure to PBDEs: Associations of PBDE Body Burdens with Food Consumption and House Dust Concentrations. Environ. Sci. Technol. 41, 1584–1589.

Yamamoto, N., Dannemiller, K.C., Bibby, K., and Peccia, J. (2014). Identification accuracy and diversity reproducibility associated with internal transcribed spacer-based fungal taxonomic library preparation. Environ. Microbiol. 16, 2764–2776.